Entrepreneurial intent and commercialization of applications on a technology platform

Fenger, Morten H J ; Frederiksen, Lars ; Juhl, Hans Jørn ; Scholderer, Joachim

Abstract: Third-party providers as entrepreneurs boost technology platforms. Yet, despite increasing interest in technological platforms, existing research offers little predictive insight into how firms can identify individuals who are likely to become entrepreneurs. We take the strategic perspective of a platform owning company, asking how to pinpoint those individuals who transition into entrepreneurship in the near future. We base our analysis on automatically registered behavioral data, such as the complete sales history of all applications related to a technological platform, and the complete history of communications in communities related to the platform. We employ logistic regression models to predict: (a) the transition from registered platform user to third-party developer (i.e. entrepreneurial intent) and (b) the launch of a first platform application (i.e. commercialization). We control for individuals' social network positions, their communication behaviors, exposure to input from other entrepreneurs (i.e social contagion), and their early adoption and lead- user traits. We show that even after inclusion of these controls volume-wise "bulk" consumption still adds significantly to the predictional power on each step towards entrepreneurship. The impact of simple measures for bulk consumption on entrepreneurship is often ignored in the entrepreneurship literature. Our study contributes to the strategic management literature on the dynamics of innovation on technological platforms, by explicitly linking the production and consumption sides of two-sided markets. It also adds to the entrepreneurship literature by showing how the entrepreneurial process manifests itself in the context of technological platforms.

Fenger, M.J.H., Frederiksen, L., Juhl, H. J. & Scholderer, J. (in press). Entrepreneurial intent and commercialization of applications on a technology platform. In *Making Organizations Meaningful: Academy of Management Proceedings 2016.* Briarcliff Manor, NY: Academy of Management.

Entrepreneurial Intent and Commercialization of Applications on a Technology Platform

Morten H. J. Fenger [1,*]

Lars Frederiksen [2]

Hans Jørn Juhl [1]

Joachim Scholderer [1,3]

[1] Department of Economics and Business Economics, Aarhus University

[2] Department of Management, Aarhus University

[3] CCRS, University of Zurich

* Corresponding author. Department of Economics and Business Economics, Aarhus University. Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. Tel.: +45 8716 6038. E-mail: mhjfenger@econ.au.dk

Entrepreneurial Intent and Commercialization of Applications on a Technology Platform: The Importance of Consumption

Abstract:

Third-party providers as entrepreneurs boost technology platforms. Yet, despite increasing interest in technological platforms, existing research offers little predictive insight into how firms can identify individuals who are likely to become entrepreneurs. We take the strategic perspective of a platform owning company, asking how to pinpoint those individuals who transition into entrepreneurship in the near future. We base our analysis on automatically registered behavioral data, such as the complete sales history of all applications related to a technological platform, and the complete history of communications in communities related to the platform. We employ logistic regression models to predict: (a) the transition from registered platform user to third-party developer (i.e. entrepreneurial intent) and (b) the launch of a first platform application (i.e. commercialization). We control for individuals' social network positions, their communication behaviors, exposure to input from other entrepreneurs (i.e social contagion), and their early adoption and lead-user traits. We show that even after inclusion of these controls volume-wise "bulk" consumption still adds significantly to the predictional power on each step towards entrepreneurship. The impact of simple measures for bulk consumption on entrepreneurship is often ignored in the entrepreneurship literature. Our study contributes to the strategic management literature on the dynamics of innovation on technological platforms, by explicitly linking the production and consumption sides of two-sided markets. It also adds to the entrepreneurship literature by showing how the entrepreneurial process manifests itself in the context of technological platforms.

INTRODUCTION

Technology platforms (henceforth: platforms) offer new and exciting ways to engage and organize entrepreneurs to boost innovation (Adner, 2012; Yoo, Boland, Lyytinen & Majchrzak, 2012). A platform is built upon an "ecosystem" owned by a company that provides the technological architecture allowing different types of users and complementary business partners to connect and benefit from the platform's base functionality (Suarez, 2012). Platforms are organised as two-sided markets where third-party providers (i.e. entrepreneurs) offer competing goods and services on one side of the market, and products compete to find end-users on the other side (Parker & Van Alstyne, 2005). The two sides of the market are intrinsically connected and influence each other's potential for success or failure (Rochet & Tirole, 2003). For platform owners, a particular challenge is the strategic management of innovation provided by new entrepreneurs (Eisenmann, Park & van Alstyne, 2006). Yet little is known about how the entrepreneurial process unfolds in such organizational settings.

Network effects occurs when more usage of products by one user increases the product's value for other users[1] (Shapiro & Varian, 2013). Platforms is an organizational form that allows for network effects and hence for reaping economies of scope of innovation (Katz & Shapiro, 1986). Third-party providers are crucial for platform-owning firms as they expand the available features and functionalities of the initial product portfolio (Boudreau, 2012) and thus generate opportunities for additional network effects for the proprietary platform owner (Mollick, 2012). Developers of third party software

[1] We refer to users as individuals since we cannot know if they use the software, even though this seems obvious in most cases.

applications (henceforth: *apps*) can make it into big business themselves[2]. We define third-party providers as individuals who, on behalf of the platform owner, develop apps targeted at end-users of the platform. Third-party providers are therefore entrepreneurs affiliated with a platform owner via arm's-length contracts (Boudreau & Lakhani, 2009) and hence third-party providers are rarely compensated directly for their development work. Still, the third-party provider chose to enter the platform owner's marketplace for apps since it typically offers greater reach of potential customers than would otherwise be available to her.[3] (Ghazawneh & Henfridsson, 2013).

To generate a successful platform strategy it is imperative for platform owners to understand the process from entrepreneurial intent to commercialization on platforms much better: namely to identify and engage potential third-party developers, who are the lifeblood sustaining a platform ecosystem. This is especially challenging since third-party developers self-select into their developer role. In this paper, we explore to what extent individual characteristics and structural conditions surrounding individuals are useful predictors of the entrepreneurial process of third-party app providers. Hereby we offer novel knowledge of the mechanisms supporting the traditional logic of entrepreneurship; that is, how individuals seek entrepreneurial rents from spotting and evaluating market opportunities, inventing and commercializing their inventions (Schumpeter, 1934). To make such inferences we capture communication- and consumption data prior to the individual's decision to register as an app developer and thus reveal entrepreneurial intent and likewise prior to the developer's commercialization of an app.

---

[2] Consider, for example, how the app "Angry Birds" expanded almost overnight into a major diversified business.
[3] In the following, we use the terms "third-party provider" and "third-party developer" interchangeably.

As such we are able to link data on the behavior of individuals on *both* sides of the market into our explanation to provide a more detailed understanding of how the supply-side of technology platforms may be influenced by demand-side activities. We empirically investigate how nascent entrepreneurs, who are notoriously difficult to identify a priori, may become informed about market opportunities and enact innovation in the context of the app economy. Being able to predict product users' entrepreneurial progress is of strategic importance for platform owners as this enables the platform owner to incentivize individuals who show developer intent, to continue their entrepreneurial process.

Via their online user community our case study firm has for more than a decade invited individuals to innovate their products, supported by a strategy of "selective revealing" (Alexy, George, and Salter, 2013). From June 2012 the company allowed individuals to develop and sell apps on their platform to complement their proprietary software products. The company chose an innovation model where individuals interested in developing apps can be granted developer license. Such license allows access to a software development kit and component library to support programming activities (von Hippel & Katz, 2002). We regard the act of obtaining a developer license as an indication of entrepreneurial intent to develop apps and thus becoming a nascent entrepreneur. The act of launching an app, in turn, we regard as an indicator of an underlying ambition to commercialize and sell apps, and thus become an entrepreneur.

We contribute by offering two pieces of explanation to the growing literature on platforms and entrepreneurs (von Hippel, 2005, Nambisan & Baron, 2012). First, we add the component of individual patterns of technology adoption of platform offerings; that is, *consumption* behavior as a "real" demand-side indicator for catalyzing or accelerating the move to become an entrepreneur. We consider this original since the literature focuses exclusively on the mere *use* of products, typically based on self-reported information on intensity and time of usage (Baldwin, Hienerth & von Hippel, 2006). We decouple our explanation of consumption as the driver of the entrepreneurial process from, for example, individuals being early adopters of technology, resonating with the idea of lead users. We do this by controlling for both the speed of adoption of new apps, and the fraction of apps sourced from other third-party developers, and hence we show that an individual's "bulk" consumption of apps constitutes a different mechanism at work than merely quick adoption. Second, a significant aspect of our study is that we only use *behavioral* data and thus do not rely on self-reported measures[4]. While behavioral data imposes some restrictions on the analysis - for example a very scarce demographic information of the analysed individuals, such data have other advantages, in particular that it comes at little or no expense for the platform owning company, and that there is no survey selection bias of whom participates in our analysis.

In sum, by employing a unique data set obtained from a platform-owning company in the music software industry we study two generations of individuals active on a platform in order to identify key mechanisms that may lead these individuals to either initiate or progress their entrepreneurial process. We demonstrate how individual bulk consumption and specific social network relations are good predictors of such behaviors.

[4] Such self-reported data as surveys, interviews, etc. naturally carry a number of bias.

THEORY

How should firms develop their platform? Theory suggests that platforms are developed as two-sided markets and that this can be done simultaneously from both sides (Parker & Van Alstyne, 2005). Still, firms face strategic challenges with regard to deciding either to promote the demand-side of the platform and its content through marketing campaigns or discounts, or to focus on involving third-party developers and thereby develop the supply-side of the platform (Gawer & Cusumano, 2012; Boudreau, 2012). If a firm follows the latter strategy, two important questions emerge: (1) how to create a suitable base of developers who can differentiate the platform products and (2) how to make sure that developers: (a) actually launch functioning inventions (e.g. apps) and (b) that such inventions meet market demand. The imperative question then is whether firms, early on, can use a simple set of leading indicators to predict which users may become developers and which of these may continue their entrepreneurial process towards commercialization. Our research aims to shed light on who these individuals are.

Existing research on strategic management of platforms has focused on a diverse range of topics, spanning entry strategies, platform quality, and consumer expectations of platform functions (Cennamo & Santalo, 2013; Zhu & Iansiti, 2012), control strategies for platform-owning companies (Boudreau, 2010), and types of competition between third-party developers (Boudreau & Jeppesen, 2014; Eisenmann, Parker & Van Alstyne, 2006). Yet, key aspects of entrepreneurship on platforms are addressed much less in this infant research domain (Gawer, 2014).

Users as innovators

Explanations of user engagement in innovation typically revolve around ability profiles (i.e., lead users; Shah & Tripsas, 2007, von Hippel, 2005) and various external stimuli such as monetary rewards, firm or peer recognition (Jeppesen & Frederiksen, 2006). A line of this research focuses on product users engaged in innovation supported by platform-related online communities. Under certain conditions, particular users' progress from sharing their advice and inventions with other users (but without remuneration) to becoming entrepreneurs and commercializing their inventions (Baldwin et al. 2006, Autio, Dahlander & Frederiksen, 2013).

Only recently are relational arguments introduced as an explanation for user innovation (Dahlander & Frederiksen, 2012). This relational view suggests that both (a) communication patterns such as how much and to how many unique others a user is connected to, and (b) structural position in the community network, have predictive value for potential engagement in innovation. This is based on two pillars of explanation: individuals' access to unique and timely information, for example, to generate market knowledge and reduce demand uncertainty and their positions in the status hierarchy of the community, for example, to enjoy benefits of legitimacy with regards to their inventions (Autio, Dahlander & Frederiksen, 2013). Still, user innovation studies on such dynamics have proven important but have largely remained descriptive since most are based on qualitative or cross-sectional data.

Entrepreneurship on platforms

Firms owning platforms must strategically manage the "input" side of their two-sided markets. We are interested in predicting the transition from being an individual consuming products on the demand-side of the market into being a developer of products, namely apps, with a (potentially evolving) ambition to launch and sell apps on the supply-side of the market. While Thornton (1999) argues that the entrepreneurship literature has little to offer regarding demand-side effects on entrepreneurial entry, White (1981) claims that the emergence of markets and thus opportunities for consumption *and* selling is an initial condition for entrepreneurship to unfold. In the setting of platforms third party providers on an entrepreneurial trajectory must move through an initial phase of entrepreneurial intent (Krueger, Reilly & Carsrud, 2000). In our research setting this is manifested and revealed by individuals registering as developers on the platform. While this is undebatably a conscious act, such early entrepreneurial intent is not necessarily a pure cognitive process of opportunity recognition and evaluation, estimating demand-side uncertainties and the trade-offs against opportunity costs (Eckhardt & Shane, 2003): It may also be driven by a combination of passion and creativity, or simply a need for creating a solution to individual problems, without ambitions to commercialize (Cardon, Wincent, Singh & Drnovsek, 2009; Shah & Tripsas, 2007). We thus acknowledge that obtaining a developer certification for an individual may only flag entrepreneurial intent and thus cannot always be understood as an act of planned behavior (Autio et al. 2013; see Ajzen, 1991, for the theory of planned behavior). The next step in the process is that the developer creates a functional product (i.e., the developer designs the functions and features of an app and writes the code), obtains approval by the platform owner's quality control unit, and finally launches the product for sale on the platform market. This market entry represents commercialization as entrepreneurial action for the developer (McMullen & Shepherd, 2006). Platform owners have an intrinsic interest in understanding the transition of product users to app entrepreneurs. In some sense platform owners thus act as venture capital businesses by creating the conditions under which entrepreneurs will thrive and strive to generate rents for both themselves and the platform owner.

Communication, network, and contagion as predictors for platform entrepreneurship

Entrepreneurship research emphasizes the importance of an individual's social relationships and network position as a catalyst for entering the process into entrepreneurship (Birley, 1986; Hoang & Antoncic, 2003; Nanda & Sørensen, 2010). Recent research shows that relationships and structural position of individuals in networks in online communities organized around particular platform products, influence their entrepreneurial behavior (von Hippel, 2007; Autio, Dahlander & Frederiksen, 2013). Social network analysis focuses on the network structure of individuals relations. Weighted and directed social networks elaborates on this structure, by quantifying the ties of the network.

We evaluate the effect of a selection of traditional social network measures, namely degree centrality, prestige and structural holes. First, individuals who are connected to a larger number of unique peers enjoy benefits of information diversity as well as quantity. Second, status hierarchies exist in online communities (Dahlander & Frederiksen, 2012) and some individuals thus enjoy particularly prestigious social positions and may therefore be preferred by other users as a source of legitimate advice on problem solving, and new ideas and visions for future product use and development (Morrison, Roberts & Midgley, 2000). Also, such individuals may, through their status position, be able to influence technological trajectories, and thereby even shape the future of demand (Autio et al., 2013). It seems to follow immediately that individuals holding such network positions may be affected by their network positions as well. Third, individuals who span otherwise unconnected individuals and thus span structural holes may obtain entrepreneurial opportunities through information arbitrage about, for example, future exchange opportunities (Burt, 2004). Additionally, they are better positioned to acquire novel information serving as input to the opportunity recognition process. Such information access also means that they may be able to better assess how ideas for potential apps are evaluated by other users (i.e., potential consumers). Each of these relational mechanisms may make it easier for individuals with such network positions to understand potential markets, recognize opportunities, and reduce demand uncertainty by this information. Jointly, these benefits facilitates venturing into entrepreneurship. These arguments are aligned with the argument that networks behave like markets where relations between individuals take the form of both pipes for information flows as well as prisms for assigning status (Podolny, 2001). Furthermore, we believe that individuals' entrepreneurial behavior are influenced by social contagion (Burt, 1987) or what economists often label: peer effects. This implies that individuals transition into entrepreneurship may be a function of their exposure to knowledge, attitudes, viewpoints of individuals who have already moved into entrepreneurship (Nanda & Sørensen, 2010). Interactions with nascent entrepreneurs may for some individuals ignite their entrepreneurial intent as such interaction may well result in specialized and domain specific knowledge- and information transfer, and potentially encouragement to venture into entrepreneurship. So far only scarce research has explored contagion effects in the domain of online communities. Also, relying on coauthor network data, Stuart and Ding (2006) argue that scientists are more likely to turn entrepreneurs if they write and publish articles with individuals who already made the transition into the commercial world. In a similar vein this suggests that if an individual in a community has more communication with others who have already exposed their entrepreneurial intent or has made it into commercialization, this may rub off and so influence the decision of our focal individual to become an entrepreneur.

Finally, social network analysis is also concerned with what information network ties may contain . However, if the network is constructed based on directed text posts, the content of ties cannot meaningfully be divided into categories. Completely disregarding the content and timing of the messages on the other hand omits a potential wealth of information (Dahlander & Piezunka 2014; Aral & Alstyne, 2011). Therefore, while acknowledging the embeddedness of individuals in social networks, information that is beyond the structural conditions of the social network can be extracted from the communication transactions of directed text messages in the community. We include such text mining measures in our analysis of progression in the entrepreneurial process[5]

Consumption as predictor for entrepreneurship

Consumption is shaped by our needs and wishes, and may have multiple facets ranging from fundamental to self-actualization (Maslow, 1943). Consumption thus reveals unobserved individual preferences, individual levels of technology adoption and engagement in a product. Still, an individual's drive for consumption is mediated by their financial situation (Slacalek, 2009). Consumption as a measure to motivate entrepreneurship has only been little explored (Viswanathan, Sridharan & Ritchie, 2010), and only as driven by demand experienced by entrepreneurs and rarely taking into consideration the development of self-efficacy for entrepreneurship obtained *via* consumption. But at times, consumption may provide more than mere utility for the

[5] These measures are explained in detail in the methods section.

purchaser, and revenue for the seller: namely opportunity to engage and to learn (Holbrook & Hirschman, 1982). Acknowledging the issues above consumption may simply be a proxy for the interest of the individual. Thus we posit that consumption serves as a predictor of entrepreneurial intent and subsequent commercialization.

Lead-users (von Hippel, 1986; Urban & von Hippel, 1988) are linked to a higher probability to become an entrepreneur (Autio et al. 2013). Lead-userness is however typically measured as self-reported statements, whereas we base our inferences on behavioral data only. A key attribute of lead-users is that they foreshadow the market and thus typically are early adopters of new products and technologies - and this can be quantified from sales transactions. Studies have shown that individuals who quickly purchase new products right after their release on the market often have an innovation disposition and may also serve as opinion leaders (Morrison, Roberts & Midgley, 2000). Early adoption reflects alertness of consumers to new technologies in the market (Rogers, 1962). Further, time to first repurchase has been linked to identifying key customers to novel products (Cardozo, Smith & Viswanathan 1988). Together, this suggests operationalizing first-mover traits for predicting entrepreneurship. Given that experiences from early adoption can be linked to entrepreneurial intent and commercialization, we speculate if other consumption patterns also may predict entrepreneurial intent and commercialization. Here we notice, that neither early-adopter traits nor lead-userness embrace the simple volume of consumption by individuals.

We coin the term *bulk consumption* to measure this volume-wise consumption by individuals. Experiences derived from consuming a product may help users develop their creativity and identify their own skills (Bem, 1972). Bandura (1977) explains how the discovery of one's own expertise may lead to increased self-efficacy. Particularly in the early stages of the entrepreneurial process, self-efficacy is shown to be a key motivational factor potentially affecting opportunity recognition and so, increasing the likelihood that entrepreneurial intent is formed and enacted (Ardichvili, Cardozo & Ray, 2003; McMullen & Shepherd, 2006). Individuals "consuming" more apps are thus slightly better equipped to detect and appreciate the potential of new apps than others (Schreier, Oberhauser & Prügl, 2007). Also in later stages of the entrepreneurial process, when an individual has already shown entrepreneurial intent, platform apps created by others may continue to serve as a source of learning. For example, ideas for apps may grow out of realisations of missing features and functionalities of offerings in the market. As marketing information about apps may downplay their drawbacks, and over-sell their merits, truly missing market offerings may only be realised by the consumers of the available offerings. Likewise, comparing one's own code to code created by others, may offer opportunities to increase one's understanding of efficient code structures in general. The same may be the case for design features. Although the process may not unfold in the manner of a coherent action plan (Autio et al., 2013), these arguments suggest that individuals with a high number of purchases of platforms apps, may be more likely to become app entrepreneurs, than individuals for whom this is not the case. We thus propose that bulk consumption may hold predictive power for entrepreneurship, even when controlling for early-adoption characteristics.

Research questions

The number of arguments offered above motivates two research questions that guide our empirical analysis. First we ask: to which degree can behavioral variables such as consumption and communication patterns predict transition from product user to become a third-party developer? That is, *do individuals who are about to become developers differ from other individuals in the community in terms of their communication activities, social network position, and consumption behavior before they display entrepreneurial intent (RQ1)?* Second, we investigate to what extent nascent entrepreneurs can be identified beforehand. That is, *do nascent entrepreneurs differ from other nascent entrepreneurs in the community before they commercialize their application in terms of their communication activities, social network position, and consumption behavior (RQ2)?*

When initiating this study we developed the basic hypothesis that consumption may relate to entrepreneurship. This was further developed as we acquired both communication and consumption data for individuals that were platform active as consumers or third party providers or both. However, as consumption itself do not just occur, but is driven by, for example, unobserved product interest and involvement, it remains difficult at this stage of research precisely to interpret what a potentially significant effect of consumption on entrepreneurship means from an organizational and an entrepreneurship perspective. Still, due to the early stage of enquiry, predicting entrepreneurship activities on the organizational form of a platform based on behavioral data alone is of such importance, that it merits early stage answers to which theoretical mechanisms are at work[6].

[6] Helfat (2007:185) maintains: "Rather than insist that empirical research always test theory, as our journals often appear to require, we can and should use empirical research to investigate phenomena that we observe in the real world."

DATA AND METHODS

Research setting

Our research setting is a firm-owned platform for music software. More than 500,000 unique user IDs are registered on the platform over the period 2002 to 2014. The firm offers a range of products for producing, processing and recording music[7]. Related to the platform is a vibrant online community that allows individuals to interact and discuss product use, future technological features and functions of products, and collaboratively solve problems related to products offered by the platform-owning firm. Individuals buy the proprietary company software (i.e., the base products of the platform) and various types of extension apps in a secure online store operated by the company. For individuals who wish to become app developers, the firm offers a software development kit and code libraries (similar to software platforms such as Apple iOS or Google Android) and access to a gated online forum for developers. Only registered developers are allowed to offer their apps on the online marketplace where they compete with other user-developed and company-developed apps. Price and consumer ratings are displayed. All apps have a file format that is reserved for products of the platform-owning firm and is incompatible with competing platforms.

The number of individuals registered as app developers reached 1,657 by end 2013. By then almost 60,000 app sales were recorded.

Close to 10,000 individuals were active writing approximately 200,000 posts in the two-year period 2012 to 2013. Of these, 4,321 individuals posted at least two messages on

[7] This research setting is also explored by, for ex., Dahlander & Frederiksen (2012) and Autio et al. (2013). These papers employ survey data, which our paper does not. Also, we add a completely new layer of information since we have gathered consumption data for all transactions between the platform owning firm and its users.

the online community in this period, with at least one message posted in 2013[8]. We restrict our focus to this most active core of platform users.[9] Entrepreneurial intent is rare among users, and commercialization is rare among nascent entrepreneurs. Therefore our data does not allow us to follow the same generation of users through entrepreneurial intent towards commercialization. Instead, we analyze two different generations of entrepreneurs. While entrepreneurship scholars may prefer to follow the same generation towards entrepreneurship, strategic management will at all points in time face different generations of individuals approaching entrepreneurship, and each transition in the process is important to understand in itself, as it increases the volume of individuals approaching entrepreneurship.

Data sources

The data for our analysis was obtained in the form of three snapshots of the platform owner's database (December 2012, December 2013, and early 2015). The database contains sales transaction data and a complete extraction of all communication on the online user forum and lists of users registered as app developers. The different data sets were linked based on user IDs and product IDs, enabling detailed data extractions on the individual user level[10], together with the construction of a social network of interactions.

[8] Writing style has been shown to change over time (e.g. Can & Patton, 2004)

[9] In robustness check we found that entrepreneurial intent was 60 times more likely to be shown by community members than other users consuming apps from the platform. This justifies the focus on community members.

[10] The data tables used in this paper were the following: *App sales list*, *Product extensions sales list*, *Extraction of all posts from user community forum*, *List of product registrations by user ID* (purchased products), *Demographic data for userIDs* (only geography variables are available), *List of users registered in a user-companies* (not dated), *List of free product extension downloads by user IDs*, *Registered app developer user IDs* (not dated), *Date where the user companies were registered*, *App product details* (Title, description and release dates), *List of user-company names with company IDs*, *List of user company IDs' launched apps*, and *Price list for apps*.

Analyses of two generations of individuals analyses were carried out to address each of our research questions. With only one context-adjusted exception, the sets of explanatory variables are identical for both analyses. Both analyses also apply a similar estimation strategy.

Analysis 1 - sample and dependent variable

Out of the 4,321 analyzed individuals, a total of 4,058 had neither yet registered themselves as app developers, nor registered a third-party company to sell apps on the platform by end-2013. One year later, 109 of these individuals had registered as app developers. Therefore, the 109 individuals represents events of *entrepreneurial intent*, and the remaining 3,949 users likewise represents non-events. As neither the registration date as a developer nor the date individuals registers a third-party app company were available due to the SQL architecture implemented by the company, the transition is inferred to take place at some point during 2014.

Analysis 2 - sample and dependent variable

We operationalize *commercialization* as an individual launching a first app. Thus, commercialization represents market entry but not the magnitude of sales success. Commercialization is conditioned on first showing entrepreneurial intent, i.e. becoming a developer given the setup by the platform owning company, since only registered app developers have access to the required software development kit. By December 2013, 219 community active nascent entrepreneurs had not yet launched their first app. One year later, 10 of these had launched their first app. These 10 new entrepreneurs represents events of commercialization, while the remaining 209 nascent entrepreneurs represents non-events.

Independent variables

All independent variables were established by December 2013, i.e. *prior* to the transition in the dependent variable in both RQ1 and RQ2. This distinction in time is important, as it otherwise would be unclear if entrepreneurship was the outcome of-, or the reason for differences in our observed independent variables.

Demographics

The platform owner have very limited access to any demographic information of community active individuals. Still their geographical location is available. While these are present from all over the world 45% of them U.S. based, and thus we add a dummy for this. One relevant but likewise unobserved predictor for entrepreneurship is technical ability (Hartog, J., Van Praag, M., & Van Der Sluis, 2010). When assessing the data, it became evident that some individuals had registered base products multiple times. We propose that the fraction between the number of product registrations (platform products) and the number of products owned (apps), may be operationalized as a proxy for the intensity of technical usage of the purchased IT products, which in turn may be interpreted as a proxy for unobserved *technical ability*.

Communication metrics

For each individual in the community the following variables describing communication were extracted: *Number of posts* by each individual in the previous two years. *Number of threads initiated* by the individual in the previous two years, operationalizing the individuals' eagerness in setting specific agendas and asking questions (Autio et al. 2013). The *Community tenure* of an individual is measured as the number of months passed since the individuals's first text message was written in the community[11]. Likewise the *number of unique contacts*. We also extracted controls related to aggregated post contents. We thus controlled for *Average length of posts*, as text volume may indicate extrovertness, which has been linked to entrepreneurial ability (Hasan & Koning, 2015). Likewise we extracted the average length of received text posts, as a measure of *Extrovert conversation partners*. We extracted a positive/negative *Sentiment coding* using the weighting scheme developed by Nielsen (2011), of each community member's aggregated posts, to control for unobserved personal characteristics. Further, we extracted a similar score for the text posts that each user received. We called this second measure *encouragement*. As a measure for the unobserved attention that each user devotes to the community, we calculated the average time to response for each user. We labelled this variable *attentiveness proxy*. This proxy was calculated as the average time between each of a user's posts and the post to which the user was responding.

Social network metrics and contagion

Social network metrics were calculated based on interactions in the online community in 2012 and 2013. Due to the architecture of the community, each post is either a new thread or a response to an existing post. This enabled us to construct a social network from the posts with individuals represented by nodes and the number of directed posts represented by weighted, directed ties[12]. Nodal *degree centrality* (Wasserman & Faust, 1994) of the individual in the community was calculated to represent informational centrality of the individual in the community. Network *prestige* was calculated as the proportion of the indegree to the out-degree of the given individual as proposed by

[11] Not limited to two years

[12] The poster was interpreted as the *sender*, and the poster of the post to which the sender was responding was interpreted as the *receiver*. A similar approach was used by Dahlander and Frederiksen (2012).

Alexander (1963). Prestige has previously been shown to influence decision-making for entrepreneurial activities (Van Praag, 1999). *Burt's constraint* measure was also calculated, as a measure of the individual's opportunity for information arbitrage by spanning structural holes in social networks (Burt, 2004). This measure is an index between zero and one, where zero means that the node is completely unconstrained and a one means that the node is completely constrained, i.e. has access to no information arbitrage in the network. Ability to span structural holes has been found to influences an individual's social capital (Walker, Kogut, & Shan, 1997) and social capital has been shown to have a positive impact on entrepreneurship (Westlund & Bolton, 2003). Therefore, we link structural hole spanning activities in the online community to the entrepreneurship process.

Another type of more specific information arbitrage may arise from whom one communicates with. We thus also extracted the specific in-degree of *Input from entrepreneurs*, motivated by the idea that it may not be the quantity of input and generic information arbitrage that drives entrepreneurship, but rather whom an individual is in contact with and hence the quality and type of information received by the focal individual. In the modeling of *entrepreneurial intent* (RQ1), the relevant input is from individuals who have already showed entrepreneurial *intent* and thus become nascent entrepreneurs in the previous period. In our modeling of commercialization (RQ2), the relevant input is from individuals already commercializing in the previous period and thus from entrepreneurs.

In some cases a given network metric may be missing for a given individual. An individual may, for example, be active in the community by starting a lot of threads, but never receive any responses and thus not be part of the network. In such cases, the missing values for the community and network metrics were set to zero in the datasets, with the exception of Burt's measure of constraint, which was coded as 1's (i.e. fully constrained). While this is not common practice, and indeed Burt (2004) defines no valid replacement of missing values, this recoding was chosen because it leads to a meaningful adjustment in the interpretation of what this variable actually measures in our context, namely into *observed* information arbitrage. It applies to all individuals that we can not rule out the possibility of individuals obtaining information arbitrage in other non-traceable ways, i.e. by reading through the +1 mio posts in the community, or by engaging in other social (platform) interactions outside the firm-hosted community.

Early adopter traits

It is known that early adopters of products may turn into entrepreneurs (Montiel & Husted, 2009). We considered four measures for this: *Product tenure* measured as months passed since the individual registered her first proprietary product from the platform owner. Product tenure is intended to measure first-mover traits on the base products. *Average time from app launch to purchase* measured as the time from each of a given individual's app purchases backwards to when the purchased app was first available to the individual. If the individual was already registered on the platform when the app was launched, then this date was the launch date. If the app was launched prior to the individual being registered on the platform, then this date in our operationalization coincides with the date that the user registered on the platform. This measure was designed to reflect the user's alertness to the app market, with lower values reflecting more alert users. Another measure for early adoption is the *Time to first repurchase*, measuring the duration between the first and the second app purchase and counting potential simultaneous app purchases as one event. A final measure for early adoption traits was defined as the *fraction of an individual's app* *purchases developed by third-party developers*. This measure was motivated by the observation, that the platform owning company itself also developed apps. However, purchases from third-party developers represented a more exploring and less conservative attitude to new product offerings.

Except for product tenure, the remaining three measures for early adoption had missing values for close to 50% of the individuals in Analysis 1, arising from multiple individuals not (yet) active in purchasing apps. No missing values were observed in Analysis 2. Missing values in Analysis 1 were imputed in the following way: *Average time from app launch to purchase* was imputed as the number of days from the earliest fingerprint of the user in the company's database until the end of the analysis period[13]. If this date was earlier than the launch of the app market by July 1.st 2012, then this market launch date was used instead of the earliest fingerprint in the calculation. The *Time to first repurchase* was calculated in the same way. The variable measuring *fraction of a user's app purchases produced by third-party sellers* was for all missing values replaced by a zero.

Bulk consumption measures

The user's consumption records were extracted from multiple transactional sales data sets, and summarized in the following variables: *Number of unique proprietary base products owned* by the individual from the platform owner. *Number of apps purchased* by the individual, representing the involvement of the individual in the app market. Where no consumption was detected for a user by a given consumption metric, we imputed a value of zero. We also control for the consumption of other extensions of the base software, the so-called "refills". We call this measure *Old-app type (dummy)*. These old apps differ from the

[13] The earliest date for product registrations and forum activity, vs. end of year 2013, respectively.

apps by only being produced by a handful of producers besides the host company, and thus represents a much less entrepreneurial market with only professional developers engaged as entrepreneurs.

Estimation strategy

Both analyses were estimated as logistic regression models.[14] In both analyses we constructed ten models by hierarchically adding the independent variables in five blocks: (1) demographics and communication characteristics, (2) social network metrics, (3) a contagion measure, (4) early adopter traits, and (5) bulk consumption metrics. The demographic variables remained in all 10 models as control variables. The independent variables were removed if they were found insignificant at the 0.1 alpha level.

Data preprocessing was performed in the software packages SAS 9.4, SAS Text Miner 13.2, and R 3.0.3. Logarithmic transformations were performed for right-skewed independent variables. Independent variables with a high proportion of zeros were coded as binary indicator variables. Univariate and bivariate descriptive statistics[15] for the original variables from Analysis 1 and 2 are reported in Table 1a and 1b respectively.

[Insert Table 1a and 1b about here]

[14] In the case of first app launches, a more specific point in time than in RQ1 could be inferred from the first app purchase date. A dynamic time-to-event analysis was therefore considered, but this idea was abandoned as it remained unclear at which point in time to start the analysis for each individual (time of "contagion").
[15] In spite of relatively high correlation coefficients, multicollinearity was not found a problem in the estimation process. Also Wasserman & Faust (1994) recommends using similar network measures although they are closely related, as they conceptually attempt to measure different structural properties of the network.

RESULTS

Table 2 reports the results of the logistic regression models predicting entrepreneurial intent in the year 2014. Model 2 shows that the log-number of posts in the community has a significant and positive effect ($p<0.001$). The effect size 0.49 is not directly interpretable, although a positive effect implies a positive effect on the probability of entrepreneurship. On the contrary, starting new threads and thus, perhaps seeking to set a new agenda is found to have a marginally significant negative effect of only half the effect size of number of threads ($p<0.1$). Our proxy for technical ability has a positive and significant effect ($p<0.001$).

After introducing the *social network metrics* in Model 3, both significant communication measures become insignificant, and both are excluded in Model 4. In Model 4, degree centrality is a positive and strongly significant predictor ($p<0.001$). The number of unique contacts remains marginally significant ($p<0.1$), and likewise the proxy for technical ability. The more refined social network metrics such as Burt's measure of constraint, prestige, and number of fora frequented, all are found insignificant.

When our measure for *social contagion* of entrepreneurial intent is included in Model 5, this parameter estimate is positive and significant, and remains so in Model 6 ($p<0.05$). However then the number of unique contacts is no longer significant, while the degree centrality remains positive and significant, although with a smaller effect size than in earlier models ($p<0.05$).

In Model 7, even when *early-adopter* traits are controlled for, the effect of social contagion remains positive and significant ($p<0.001$), while the effect of degree centrality is dominated.. In the trimmed Model 8 the proportion of apps purchased from third-party developers remains a positive and significant predictor ($p<0.001$).

When the *bulk consumption* measures are added to the model (i.e. Model 9) our technical ability proxy becomes insignificant. Bulk consumption also dominates the previously significant early adopter trait reported above. In the reduced Model 10, we observe a positive and significant effect of increased base product usage ($p<0.01$), and further a positive effect of increased app consumption ($p<0.001$). The positive effect of social contagion likewise remains in our final model ($p<0.001$).

The Max-rescaled R-squared measure of the model suggests that the final model predict 12.5% of the variation in who among the community members that shows entrepreneurial intent within the next year. To further quantify how well our models identify future nascent entrepreneurs, we also calculate the area under the ROC curve (onwards: *ROC index*) see Hanley & McNeil, 1982, for an introduction). The ROC curve plots the sensitivity[16] against 1-specificity[17], and the ROC index ranges between 0 and 1, where a completely random prediction of events will result in a ROC index of 0.5. The ROC index of Model 10 is 0.762. Both the Max-rescaled R-squared measure and the ROC index increase along models 1 to 10, and reaches their highest levels in Model 10.

[Insert Table 2 about here]

The results from Analysis 2 modeling commercialization are presented in Table 3. Our analysis follows the same modelling approach as analysis 1. In Model 1 and 2 we observe a significant and positive effect of our measure for *encouragement* ($p<0.01$). This effect remains positive and statistically significant through all subsequent models ($p<0.01$).

[16] Correctly classified events
[17] Wrongly classified non-events, i.e. false positives

When we introduce *social network* metrics in Model 3, only forum tenure is found to have a marginally significant negative effect (p<0.1). Our measure for *social contagion* does not add significantly to the prediction when introduced in Model 5. In Model 7 the *early adopter* traits were added, and in the reduced Model 8, the time to first re-purchase of apps was found to be a positive and significant predictor (p<0.01), while forum tenure becomes insignificant. When *bulk consumption* is controlled for in Model 9, all three measures are found to add significantly to the prediction: The number of base products registered has a positive effect (p<0.05), old app-type usage has a negative effect (p<0.01), and volume of app consumption has a positive effect (p<0.05). The early-adopter trait *Time to first repurchase of app* remains significant in the final model as well (p<0.05). The Max-rescaled R-squared measure suggests 41% of the variation is explained by Model 10, where the ROC index reaches 0.899. [Insert Table 3 about here]

Social network metrics were not significant predictors for commercialization in Model 3. Note however that nascent entrepreneurs already differs from regular individuals in both communication activity and social network positions as summarized in Table 4.[18]

[Insert Table 4 about here]

We summarize our results across the two models in Table 5. It becomes clear that different predictors are important at different points progressing through the entrepreneurial process, whereas bulk consumption measures are positive predictors in both analysis.

[Insert Table 5 about here]

[18] Further, the 4,321 individuals in our sample are specifically selected based on communication activity, which also forms the basis for the social network

Robustness Checks

In both analysis, we also estimated models 2, 4, 6, and 8 by including also insignificant variables from the previous models. This did not change the results reported in table 2 and 3. Significant correlations between the independent variables in both analysis led us further to inspect for multicollinearity. Although this is not necessarily a problem (see O'brien, 2007) we calculated the variance inflation factors (VIF). These all are below 10.

The non-negligible fraction of missing values (close to 50%) of multiple first mover traits in Analysis 1 could lead to question our findings of the effect of bulk consumption. We thus replicated our results without imputation of missing values, with the only significant change being that the share of apps sourced from third-party developers would no longer be significant in Analysis 1, model 8. Our findings of the effect of bulk consumption are thus unaffected by this imputation.

We cross-validated our final models from Analysis 1 and 2 on hold-out data with a 10-folds cross validation procedure (Friedman, Hastie & Tibshirani, 2001). Applying a cut-off probability of 0.05, cross validation sensitivity in Analysis 1 reached 44.0%, with a false positive rate of 12.1%. The ROC index reached 0.7618. These values are very close to the figures reported in Table 2.

In Analysis 2, cross validation sensitivity reached 80.0%, with a false positive rate of 15.4%, and a ROC index of 0.8127 - again close to figures reported in Table 3. In summary, this is cross validated evidence for out-of-sample generalizability of our results. Thus, while both Analysis 1 and in particular Analysis 2 classifies rare events on a comparatively large set of explanatory variables, the final models are successful in classifying hold-out observations.

DISCUSSION AND CONCLUSION

For platform owners it is vital to be able to develop the supply-side of their innovation ecosystem. Our empirical analysis offers identification of those individuals who are likely to move from platform users to become third party developers in the following time period as well as for predicting who are likely to progress into commercializing an app.

In our results for predicting entrepreneurial intent contagion dominates other social network measures. Contagion is a focused social network measure, where some peers have more impact than others. Contagion of entrepreneurship through peer effects has been shown in both high-tech environments (Zhang, 2003) and employing register data (Nanda & Sørensen, 2008). Interestingly, only our bulk consumption measures adds significantly to the prediction of entrepreneurial intent, once contagion is controlled for. It should be kept in mind, that bulk consumption in turn may be driven by unobserved product involvement, which previously has been linked to entrepreneurship (Atuahene-Gima & Ko, 2001). Thus, while we theoretically speculate that bulk consumption may reinforce such product involvement, our contribution from an organizational perspective is unaffected by this distinction: Bulk consumption can be extracted from sales records alone, and is found as the most persistent significant predictor for both entrepreneurial intent and commercialization.

Interestingly, our results for prediction of first app commercialization indicate that once a nascent entrepreneur, social network position does not add further to the likelihood of taking entrepreneurial action. However, controlling for received encouragement in the form of positive talking, as evaluated by sentiment coding of received text posts, is a strong positive and significant predictor. While for driving entrepreneurial *intent* it was found to matter *whom* among the peers that was taking to alter, in RQ2 it was found to matter *what* peers wrote to alter to effect behavior. That encouragement may spur entrepreneurship is hardly novel but existing research tend to focus on encouragement of entrepreneurship by policy (Baumol, 1968; Lerner, 2010) or venture capital, while research on encouragement as the *content* of peer influence to a lesser degree has been empirically associated with commercialization activity. The encouragement they receive may reduce demand uncertainties, and provide them with a feeling of ability to provide useful products to other users (Wasko & Faraj, 2000; Autio et al. 2013). This response from a potential market of users may in turn drive their confidence in being successful in their entrepreneurial endeavor. Controlling for encouragement in the prediction of commercialization, our bulk consumption measures still adds significantly. Hereby we complement Mollick's (2012) study, with the refinement that while intrinsic motivations such as "encouragement" may hold elements of encouragement towards commercialization as well.

Implications for research

We contribute to three strands of research. First, we add to prior studies that demonstrate that user communities constitute an important determinant in explaining entrepreneurship (Autio et al. 2013). We connect this literature to the current discussion about organization of innovation via platforms (Gawer, 2014; Boudreau, 2012). Second, our findings introduce a new type of explanation that should be included for obtaining a better understanding of the transition process into entrepreneurship on platforms, namely the consumption history of individuals. By highlighting how consumption patterns play a role in predicting entrepreneurship we believe we are opening an original agenda for future research. This agenda is interested in explanations from the 'real' demand-side and thus to include insights on innovation arriving from the field of marketing (Hauser, Tellis & Griffin, 2006). Third, the attention to purchase history as extracted from sales transactions, offers a different approach to studies of entrepreneurship than self-reported product use, product related abilities, or how certain stimuli affect motivation for invention (von Hippel, 2005; Jeppesen & Frederiksen, 2006; Shah, 2006). Whereas research on entrepreneurship yields important insights into the organization of entrepreneurship, too often the focus has ignored behavioral measures. We believe our study serves as an inspiration for exploiting big data in this regard[19].

In sum, we offer two insights to entrepreneurship studies. First, research tends to theorize about opportunities themselves (Shane & Venkataraman, 2000) rather than examine the conditions that prompt individuals to perceive opportunities, evaluate them, and act upon them (Choi & Shepherd, 2004). This has led the conceptual lens to focus on the entrepreneurial process (McMullen & Shepherd, 2006). We explore empirically how the entrepreneurial process unfolds and, in particular, we demonstrate via our two-stage study that only some parts of our behavioral explanations are useful for predicting entrepreneurial intent and commercialization. Thus, while we caution to think that our explanations are equally useful for predicting the entrepreneurial process at each stage, we still maintain that consumption measures are applicable for making predictions on both assessed entrepreneurial stages. That different mechanisms are found to contribute significantly at different stages in the entrepreneurial process, supports the view that entrepreneurship is perhaps not to understand as one homogenous strive towards commercialization of an invention. Rather entrepreneurship is usefully thought of as a

---

[19] We acknowledge that it is challenging to obtain data that allows for studies of how users over time consume, communicate as well as produce innovation but still encourage researchers to explore this new trajectory.

series of hierarchically ordered different stages, where different factors determine the individual's choice for progressing onto pursuing opportunities at the next stage (McMullen & Shepherd, 2006).

Second, research shows that social relationships and network position influence individuals' decision to become an entrepreneur (Birley, 1986; Nanda & Sørensen, 2010). Yet, precisely how such mechanisms work in an online setting is still less evident. We demonstrate that mechanisms such as contagion and encouragement from already established entrepreneurs may predict entrepreneurship observed in the influenced individuals.

Limitations and future research agendas

Though we advocate the similarity between platforms where individuals can become contributors (Autio et al. 2013) we acknowledge that the generalizability of our results may be discussed as we derived our findings from only one such organizational form. Thus, despite the inherent difficulty in gathering data for large-scale empirical investigations in similar settings as ours, we encourage replication studies of drivers for the entrepreneurship process in various platform setting. Also, we encourage additional research on the effect of consumption on predicting entrepreneurship for other product types than apps, and in different organizational settings.

Conclusion

Our study breaks new ground for the understanding of the entrepreneurial process in a platform organization. Our study informs strategic management of platform owners. By an empirical study we emphasize how behavioral data, accessible and free of charge for the platform owner, provide the basis for prediction of progression through different stages of entrepreneurship. Also, we contribute by offering an original type of explanation for this entrepreneurial process by emphasizing how the consumption histories of individuals impact their future transition into entrepreneurship. This explanation complements other explanations for entrepreneurship on platforms that we offer, namely that entrepreneurial intent is also more frequent among individuals exposed to direct input from other nascent entrepreneurs in the online community. Furthermore, commercialization is more likely for nascent entrepreneurs who are encouraged to take this action by their peers. As the number of platforms rapidly increases and gain significance in the economy, our study enriches the literature on organizing innovation by offering insights on the dynamics of entrepreneurship in such ecosystems, and how consumption may preseed entrepreneurship.

REFERENCES

Adner, R. (2012). *The wide lens: A new strategy for innovation*. Penguin UK.

Ajzen, I. (1991). The theory of planned behavior. Organizational behavior and human decision processes, 50(2), 179-211.

Alexy, O., George, G., & Salter, A. J. (2013). Cui bono? The selective revealing of knowledge and its implications for innovative activity. *Academy of Management Review*, 38(2), 270-291.

Anderson, J. (1983). LIX and RIX: Variations on a little-known readability index. *Journal of Reading*, 490-496.

Aral, S., & Van Alstyne, M. (2011). The diversity-bandwidth trade-off1. *American Journal of Sociology*, 117(1), 90-171.

Ardichvili, A., Cardozo, R., & Ray, S. (2003). A theory of entrepreneurial opportunity identification and development. *Journal of Business venturing*, 18(1), 105-123.

Atuahene-Gima, K., & Ko, A. (2001). An empirical investigation of the effect of market orientation and entrepreneurship orientation alignment on product innovation. *Organization science*, 12(1), 54-74.

Autio, E., Dahlander, L., & Frederiksen, L. (2013). Information exposure, opportunity evaluation and entrepreneurial action: An empirical investigation of an online user community. *Academy of Management Journal*, 56(5), 1348-1371.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191.

Bem, D. J. (1972). Self-perception theory. *Advances in experimental social psychology*, 6, 1-62.

Birley, S. (1986). The role of networks in the entrepreneurial process. *Journal of Business Venturing*, 1(1), 107-117.

Baldwin, C., Hienerth, C., & Von Hippel, E. (2006). How user innovations become commercial products: A theoretical investigation and case study. Research policy, 35(9), 1291-1313.

Boudreau, K. (2010). Open platform strategies and innovation: Granting access vs. devolving control. *Management Science*, *56*(10), 1849-1872.

Boudreau, K. J. (2012). Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization Science*, 23(5), 1409-1427.

Boudreau, K. J., & Jeppesen, L. B. (2014). Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal*.

Boudreau, K. J. & Lakhani, K. R. 2009. How to Manage Outside Innovation. *MIT Sloan Management Review*, 50(4): 68-76

Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349-399.

Can, F., & Patton, J. M.. (2004). Change of Writing Style with Time. *Computers and the Humanities*, *38*(1), 61–82.

Cardozo, R. N., Smith, D. K., & Viswanathan, M. (1988). Identifying key customers for novel industrial products. *Journal of Product Innovation Management*, *5*(2), 102-113.

Cennamo, C. & Santalo, J. (2013). Platform competition: Strategic trade-offs in platform markets. *Strategic Management Journal*, 34(11): 1331-1350.

Choi, Y. R., & Shepherd, D. A. (2004). Entrepreneurs' decisions to exploit opportunities. *Journal of Management*, 30(3), 377-395.

Dahlander, L., & Frederiksen, L. (2012). The core and cosmopolitans: A relational view of innovation in user communities. *Organization Science*, 23(4), 988-1007.

Dahlander, L., & Piezunka, H. (2014). Open to suggestions: How organizations elicit suggestions through proactive and reactive attention. *Research Policy*,*43*(5), 812-827.

Eisenmann, T., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard Business Review*, 84(10), 92.

Eckhardt, J. T., & Shane, S. A. (2003). Opportunities and entrepreneurship. *Journal of Management*, 29(3), 333-349.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1). Springer, Berlin: Springer series in statistics.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, 43(7), 1239-1249.

Gawer, A., Cusumano, M. A., & Strategy, D. S. (2012). How companies become platform leaders. *MIT/Sloan Management Review*, 49.
Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third□ party development: the boundary resources model. *Information Systems Journal*, 23(2), 173-192.
Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1), 29-36.
Hasan, S., Stanford, G. S. B., & Koning, R. (2015). Conversational Peers and Idea Generation: Evidence from a Field Experiment. *Working paper - Stanford.edu.*
Hauser, J., Tellis, G. J., & Griffin, A. (2006). Research on innovation: A review and agenda for marketing science. *Marketing Science*, *25*(6), 687-717.
Helfat, C. E. (2007). Stylized facts, empirical research and theory development in management. *Strategic Organization*, 5(2), 185-192.
Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, 132-140.
Hoang, H., & Antoncic, B. (2003). Network-based research in entrepreneurship: A critical review. *Journal of Business Venturing*, 18(2), 165-187.
Jeppesen, L. B., & Frederiksen, L. (2006). Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science*, 17(1), 45-63.
Katz, M. L., & Shapiro, C. (1986). Technology adoption in the presence of network externalities. *The Journal of Political Economy*, 822-841.
Krueger, N. F., Reilly, M. D., & Carsrud, A. L. (2000). Competing models of entrepreneurial intentions. *Journal of Business Venturing*, 15(5), 411-432.
Lerner, J. (2010). The future of public efforts to boost entrepreneurship and venture capital. *Small Business Economics*, 35(3), 255-264.
Maslow, A. H. (1943). A theory of human motivation. *Psychological review*, 50(4), 370.
McMullen, J. S., & Shepherd, D. A. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review*, 31(1), 132-152.
Mollick, E.R. (2012). Filthy Lucre: What Motivates the Commercialization of Innovations? (working paper, SSRN).
Montiel, I., & Husted, B. W. (2009). The adoption of voluntary environmental management programs in Mexico: First movers as institutional entrepreneurs. *Journal of Business Ethics*, 88(2), 349-363.
Morrison, P. D., Roberts, J. H., & Midgley, D. F. (2000). Opinion leadership amongst leading edge users. *Australasian Marketing Journal*, 8(1), 5-14.
Nambisan, S., & Baron, R. A. (2013). Entrepreneurship in Innovation Ecosystems: Entrepreneurs' Self□ Regulatory Processes and Their Implications for New Venture Success. *Entrepreneurship Theory and Practice*, 37(5), 1071-1097.
Nanda, R., & Sørensen, J. B. (2010). Workplace peers and entrepreneurship. *Management Science*, 56(7), 1116-1126.
Nanda, R., & Sørensen, J. B. (2008). Peer effects and entrepreneurship. Harvard Business School.
Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *arXiv preprint arXiv*:1103.2903.
O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673-690.
Parker, G. G., & Van Alstyne, M. W. (2005). Two-sided network effects: A theory of information product design. *Management Science*, 51(10), 1494-1504.
Podolny, J. M. (2001). Networks as the Pipes and Prisms of the Market 1. *American Journal of Sociology*, 107(1), 33-60.
R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990-1029.
Rogers EM. (1962). *Diffusion of Innovations*. Free Press of Glencoe, Macmillan Company.
SAS Institute. (2014). SAS software. Copyright, SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.
Schreier, M., Oberhauser, S. & Prügl, R.W. (2007) Lead users and the adoption and diffusion of new products: Insights from two extreme sports communities. *Marketing Letters*, 18 (1-2). pp. 15-30.

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle* (Vol. 55). Transaction publishers.
Shah, S. K. (2006). Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, 52(7), 1000-1014.
Shah, S. K., & Tripsas, M. (2007). The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal*, 1(1□2), 123-140.
Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review*, 25(1), 217-226.
Shapiro, C., & Varian, H. R. (2013). Information rules: a strategic guide to the network economy. *Harvard Business Press.*
Suarez, F. F. (2012). Dethroning an established platform. *MIT Sloan Management Review*, 53(4).
Thornton, P. H. (1999). The sociology of entrepreneurship. *Annual Review of Sociology*, 19-46.
Urban, G. L., & Von Hippel, E. (1988). Lead user analyses for the development of new industrial products. *Management science*, *34*(5), 569-582.
Viswanathan, M., Sridharan, S., & Ritchie, R. (2010). Understanding consumption and entrepreneurship in subsistence marketplaces. *Journal of Business Research*, 63(6), 570-581.
von Hippel, E. (1986). Lead users: a source of novel product concepts. *Management Science*, 32(7), 791-805.
von Hippel, E. (2005). *Democratizing innovation*. Cambridge, Massachusetts.
von Hippel, E. (2007). Horizontal innovation networks—by and for users. *Industrial and Corporate Change*, 16(2), 293-315.
von Hippel, E., & Katz, R. (2002). Shifting innovation to users via toolkits. *Management Science*, 48(7), 821-833.
Walker, G., Kogut, B., & Shan, W. (1997). Social capital, structural holes and the formation of an industry network. *Organization science*, *8*(2), 109-125.
Wasko, M. M., & Faraj, S. (2000). "It is what one does": why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems*, 9(2), 155-173.
Wasserman, S. & Faust, K. (1994). *Social network analysis: Methods and applications*(Vol. 8). Cambridge university press.
Westlund, H., & Bolton, R. (2003). Local social capital and entrepreneurship.*Small business economics*, *21*(2), 77-113.
White, H. C. (1981). Where do markets come from? *Advances in Strategic Management*, 17(2), 323-350.
Yoo, Y., Boland Jr, R. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, 23(5), 1398-1408.
Zhang, J. (2003). Growing Silicon Valley on a landscape: an agent-based approach to high-tech industrial clusters. *Journal of Evolutionary Economics*,13(5), 529-548.
Zhu, F., & Iansiti, M. (2012). Entry into platform□based markets. *Strategic Management Journal*, 33(1), 88-106.

APPENDIX

Tables

Table 1a. Correlations and descriptive statistics

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Nascent entrepreneur	1														
2	US residence (dummy)	-0,023	1													
3	Technical ability proxy	0,094***	-0,008	1												
4	More than one post (dummy)	0,063***	-0,011	0,127***	1											
5	Log (No. of threads started + 1)	0,072***	0,033*	0,095***	0,34***	1										
6	Average length of posts	-,082***	-0,004	-0,095***	0,202***	-0,371***	1									
7	Extrovert conversation partners	0,02	-0,037*	0,106***	0,441***	-0,024	0,461***	1								
8	Sentiment coding (positive)	0,045**	-0,013	0,075***	0,244***	0,31***	-0,142***	0,157***	1							
9	Encouragement (inbound sentiment coding)	0,081***	0,01	0,119***	0,273***	0,322***	-0,224***	0,25***	0,546***	1						
10	Attentiveness proxy	-0,004	0,026	-0,051***	-0,172***	-0,102***	0,033*	-0,174***	-0,101***	-0,115***	1					
11	Forum tenure	0,047**	0,041**	0,055***	0,153***	0,087***	-0,089***	0,09***	0,109***	0,132***	-0,079***	1				
12	Log (No. of unique contacts + 1)	0,104***	-0,044**	0,18***	0,195***	0,176***	-0,298***	0,091***	0,234***	0,268***	-0,094***	0,179***	1			
13	No. of fora frequented	0,125***	-0,035*	0,226***	0,326***	0,582***	-0,429***	0,052***	0,339***	0,383***	-0,06***	0,186***	0,532***	1		
14	Log (Degree centrality)	0,132***	-0,012	0,219***	0,51***	0,645***	-0,455***	0,281***	0,454***	0,533***	-0,227***	0,249***	0,483***	0,72***	1	
15	Prestige	-0,025	0,052***	0,014	-0,042**	0,131***	0,045**	0,044**	-0,12***	-0,077***	0,064***	-0,054***	-0,148***	-0,074***	-0,003	1
16	Burt's measure of constraint	-,086***	0,014	-0,202***	-0,583***	-0,372***	0,051**	-0,496***	-0,349***	-0,388***	0,22***	-0,214***	-0,369***	-0,478***	-0,771***	-0,116***
17	Log (In-degree from entrepreneurs +1)	0,137***	0	0,12***	0,244***	0,44***	-0,447***	0,011	0,284***	0,321***	-0,107***	0,326***	0,397***	0,545***	0,655***	-0,075***
18	Product tenure (base product)	-0,012	0,001	-0,384***	0,043**	-0,037*	0,011	0,019	0,001	0,01	-0,011	0,394***	0	-0,017	0,02	-0,021
19	Avg. time to purchase of apps	0,015	-0,036	-0,008	-0,032	-0,009	0,055*	-0,005	0	-0,035	0,011	0,016	-0,051*	-0,022	-0,074**	0,023
20	Time to first repurchase of apps	-0,032	0,027	-0,09***	-0,065**	-0,086***	0,12***	0,008	-0,056*	-0,103***	0,067**	0,035	-0,069**	-0,122***	-0,142***	0,073**
21	Pct. apps from third-party developers	0,031	-0,001	0,013	0,084***	0,081***	-0,053*	0,025	0,043	0,041	-0,01	-0,013	0,058*	0,11***	0,114***	-0,038
22	No. of base product purchased	0,124***	0,001	0,806***	0,153***	0,14***	-0,127***	0,111***	0,109***	0,148***	-0,058***	0,09***	0,236***	0,311***	0,274***	-0,025
23	Old app-type (dummy)	0,031*	0,031*	0,247***	0,112***	0,126***	-0,11***	0,059***	0,121***	0,149***	-0,048**	0,118***	0,119***	0,198***	0,226***	-0,011
24	Log (No. of app purchases + 1)	0,134***	-,092***	0,289***	0,198***	0,223***	-0,221***	0,1***	0,221***	0,253***	-0,094***	0,15***	0,292***	0,369***	0,425***	-0,081***
	N	4134	4134	4134	4134	4134	4058	4134	4058	4134	4134	4134	4134	4134	4134	4134
	Mean	0,03	0,44	0,5	0,75	1,08	714,8	480,43	0,14	0,09	0,26	33,21	0,63	1,93	2,35	0,55
	Std Dev	0,16	0,5	0,36	0,43	0,83	242,34	381,84	0,19	0,16	0,83	36,01	0,96	1,28	1,39	0,22
	Skewness	5,91	0,26	0,19	-1,16	1,21	-1,18	-0,09	0,9	1,23	7,41	2,44	1,63	1,95	1	-0,24
	Kurtosis	33	-1,94	0,52	-0,66	2,01	0,85	-1,66	1,61	2,69	73,47	19,65	2,07	4,86	0,72	1,14
	Minimum	0	0	0	0	0	6,36	0	-0,48	-0,55	0	0,06	0	1	0	0
	Maximum	1	1	2,5	1	5,37	1000	1000	1,16	1,03	11,88	635,87	4,89	10	7,2	1

	Variable	16	17	18	19	20	21	22	23	24
16	Burt's measure of constraint	1								
17	Log (In-degree from entrepreneurs +1)	-,393***	1							
18	Product tenure (base product)	-0,028	0,103***	1						
19	Avg. time to purchase of apps	0,052*	-0,076**	0,001	1					
20	Time to first repurchase of apps	0,091***	-0,061**	0,102***	-0,042	1				
21	Pct. apps from third-party developers	-,079***	0,089***	-0,025	-0,136***	-0,137***	1			
22	No. of base product purchased	-,238***	0,169***	-0,297***	0,012	-0,101***	0,021	1		
23	Old app-type (dummy)	-,193***	0,108***	-0,026	-0,022	-0,103***	-0,007	0,261***	1	
24	Log (No. of app purchases + 1)	-,333***	0,281***	-0,104***	-0,13***	-0,567***	0,2***	0,351***	0,35***	1
	N	4134	4134	4134	1821	1821	1821	4134	4134	4134
	Mean	0,35	0,65	13,76	82,49	83,81	0,85	1,49	0,5	0,81
	Std Dev	0,31	1,18	20	75,13	103,31	0,26	1,32	0,5	1,09
	Skewness	1,06	2,02	3,99	0,96	1,41	-2,21	0,68	-0,01	1,15
	Kurtosis	-0,04	3,58	16,79	1,62	0,81	4,2	-0,04	-2	0,21
	Minimum	0,01	0	0,03	-274,5	1	0	0	0	0
	Maximum	1	6,55	139,58	378	383	1	10	1	4,84

***p<0.001; **p<0.01; *p<0.05.

Table 1b. Correlations and descriptive statistics

	Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Nascent entrepreneur	1														
2	US residence (dummy)	0,085	1													
3	Technical ability proxy	0,082	0,167*	1												
4	More than one post (dummy)	0,002	0,042	0,159*	1											
5	Log (No. of threads started + 1)	0,087	0,1	0,223***	0,304***	1										
6	Average length of posts	-0,068	-0,058	-0,185**	-0,035	-0,703***	1									
7	Extrovert conversation partners	-0,051	0	-0,057	0,291***	-0,365***	0,679***	1								
8	Sentiment coding (positive)	0,115	0,074	0,218**	0,238***	0,404***	-0,394***	-0,113	1							
9	Encouragement (inbound sentiment coding)	0,25***	0,088	0,223***	0,273***	0,44***	-0,397***	-0,069	0,534***	1						
10	Attentiveness proxy	-0,025	-0,053	-0,124	-0,232***	-0,226***	0,109	-0,074	-0,261***	-0,248***	1					
11	Forum tenure	-0,063	0,086	0,166*	0,094	0,175**	-0,179**	-0,102	0,217**	0,181**	-0,067	1				
12	Log (No. of unique contacts + 1)	0,025	0,042	0,234***	0,176**	0,367***	-0,483***	-0,199**	0,325***	0,302***	-0,216**	0,212**	1			
13	No. of fora frequented	0,038	0,178**	0,32***	0,274***	0,744***	-0,697***	-0,375***	0,417***	0,468***	-0,188**	0,274***	0,598***	1		
14	Log (Degree centrality)	0,065	0,135*	0,292***	0,453***	0,783***	-0,762***	-0,228***	0,504***	0,537***	-0,308***	0,267***	0,558***	0,799***	1	
15	Prestige	-0,005	0,103	0,011	-0,131	-0,015	0,161*	0,087	-0,119	-0,108	0,153*	-0,023	-0,147*	-0,134*	-0,158*	1
16	Burt's measure of constraint	-0,017	-0,089	-0,263***	-0,546***	-0,452***	0,319***	-0,18**	-0,375***	-0,414***	0,302***	-0,201**	-0,408***	-0,462***	-0,743***	0,068
17	Log (In-degree from entrepreneurs +1)	0,074	0,126	0,221***	0,219***	0,593***	-0,689***	-0,384***	0,392***	0,364***	-0,156*	0,236***	0,456***	0,644***	0,744***	-0,146*
18	Product tenure (base product)	-0,04	0,014	-0,462***	-0,06	-0,063	0,098	-0,005	-0,002	-0,051	0,032	0,211**	-0,104	-0,069	-0,076	0,076
19	Avg. time to purchase of apps	0,131	0,022	-0,176**	-0,041	-0,246***	0,219**	0,013	-0,196**	-0,161*	0,084	-0,04	-0,206**	-0,169*	-0,276***	-0,03
20	Time to first repurchase of apps	0,137*	0,016	-0,128	-0,082	-0,268***	0,196**	-0,006	-0,202**	-0,176**	0,127	-0,031	-0,166*	-0,178**	-0,292***	0,013
21	Pct. apps from third-party developers	-0,115	-0,002	0,139*	0,093	0,267***	-0,231***	-0,017	0,199**	0,175**	-0,099	0,064	0,169*	0,204**	0,293***	0,021
22	No. of base product purchased	0,127	0,14*	0,803***	0,136*	0,243***	-0,178**	-0,033	0,16*	0,224***	-0,126	0,084	0,213**	0,346***	0,288***	-0,029
23	Old app-type (dummy)	-0,074	0,077	0,155*	-0,027	0,111	-0,083	-0,058	0,065	0,09	0,069	0,057	0,107	0,08	0,092	0,107
24	Log (No. of app purchases + 1)	-0,064	-0,043	0,151*	0,117	0,324***	-0,271***	-0,075	0,196**	0,224***	-0,108	0,049	0,185**	0,23***	0,337***	-0,045
	N	223	223	223	223	223	219	223	219	223	223	223	223	223	223	223
	Mean	0,04	0,41	0,72	0,9	1,55	611,72	543,32	0,19	0,17	0,25	52,87	1,16	3,19	3,42	0,53
	Std Dev	0,21	0,49	0,3	0,3	1,18	335,05	372,83	0,2	0,18	0,75	40,97	1,24	2,34	1,68	0,16
	Skewness	4,43	0,38	-0,34	-2,63	0,78	-0,54	-0,24	0,43	0,66	5,1	0,68	0,86	1,61	0,31	0,15
	Kurtosis	17,77	-1,88	2,38	4,95	0,18	-1,17	-1,57	0,59	0,22	30,24	-0,84	-0,31	2,76	-0,68	3,88
	Minimum	0	0	0	0	0	6,19	0	-0,28	-0,24	0	0,06	0	1	0	0
	Maximum	1	1	2	1	4,94	1000	1000	0,92	0,86	6,22	139,58	4,36	12	7,34	1

	Variable	16	17	18	19	20	21	22	23	24
16	Burt's measure of constraint	1								
17	Log (In-degree from entrepreneurs +1)	-0,419***	1							
18	Product tenure (base product)	0,015	-0,005	1						
19	Avg. time to purchase of apps	0,276***	-0,106	0,192**	1					
20	Time to first repurchase of apps	0,3***	-0,064	0,186**	0,921***	1				
21	Pct. apps from third-party developers	-0,297***	0,124	-0,14*	-0,906***	-0,869***	1			
22	No. of base product purchased	-0,25***	0,216**	-0,335***	-0,2**	-0,161*	0,187**	1		
23	Old app-type (dummy)	-0,149*	0,029	-0,068	-0,352***	-0,326***	0,292***	0,13	1	
24	Log (No. of app purchases + 1)	-0,302***	0,124	-0,151*	-0,82***	-0,851***	0,797***	0,193**	0,395***	1
	N	223	223	223	223	223	223	223	223	223
	Mean	0,19	1,03	12,45	218,18	197,07	0,61	2,79	0,57	1,48
	Std Dev	0,25	1,39	16,64	217,29	230,81	0,44	1,55	0,5	1,19
	Skewness	2,1	1,2	5,16	0,66	0,69	-0,56	0,03	-0,28	0,2
	Kurtosis	3,93	0,34	27,46	-1,37	-1,38	-1,56	-0,08	-1,94	-0,96
	Minimum	0,01	0	0,42	-29	1	0	0	0	0
	Maximum	1	5,39	123,74	537	537	1	8	1	4,22

***p<0.001; **p<0.01; *p<0.05.

Table 2: Logistic regression on entrepreneurial intent

Predictor variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
COMMUNICATION										
Demographics										
US residence (dummy)	-0,2268 (0,2051)	-0,2268 (0,2046)	-0,2039 (0,2057)	-0,212 (0,205)	-0,2181 (0,2055)	-0,2385 (0,2048)	-0,168 (0,2058)	-0,1706 (0,2055)	-0,1588 (0,2071)	-0,1773 (0,207)
Technical ability proxy	1,2448*** (0,2723)	1,2599*** (0,27)	1,2131*** (0,2787)	1,2225*** (0,275)	1,194*** (0,2734)	1,2253*** (0,2703)	1,2729*** (0,3074)	1,1844*** (0,288)	0,4159 (0,4736)	0,394 (0,4668)
Communication metrics										
Log (No. of posts)	0,5012*** (0,1271)	0,4923*** (0,0832)	0,0387 (0,4278)							
Log (No. of threads started + 1)	-0,2224 (0,1394)	-0,2515' (0,1381)	-0,1379 (0,1624)							
Average length of posts	-0,00006 (0,000702)									
Extrovert conversation partners	0,00039 (0,000422)									
Sentiment coding (positive tone)	-0,5517 (0,5794)									
Encouragement (inbound sentiment coding)	0,7211 (0,6591)									
Attentiveness proxy	0,1659' (0,0991)									
Social network metrics										
Forum tenure			0,00228 (0,00253)							
Log (No. of unique contacts + 1)			0,1057 (0,1061)	0,1631' (0,096)	0,1342 (0,0985)					
No. of fora frequented			0,0554 (0,0943)							
Log (Degree centrality)			0,2772 (0,4647)	0,3384*** (0,0698)	0,203* (0,0945)	0,244** (0,0894)	0,1203 (0,0927)			
Prestige			-0,4928 (0,7376)							
Burt's measure of constraint			-0,4041 (0,8034)							
Contagion										
Log (In-degree from nascent entrepreneurs +1)					0,1973* (0,089)	0,2131* (0,088)	0,2296* (0,0893)	0,3308*** (0,0592)	0,2935*** (0,0607)	0,3028*** (0,0603)
CONSUMPTION										
Early adopter metrics										
Product tenure (base product)							0,00723 (0,00605)			
Avg. time to purchase of apps							0,000867 (0,0013)			
Time to first repurchase of apps							-0,00156 (0,00112)			
Pct. apps from third-party developers							0,7673 (0,5594)	1,1316*** (0,2431)	0,5729 (0,3498)	
Bulk consumption metrics										
No. of base product purchased									0,2908** (0,1063)	0,291** (0,1047)
Old app-type consumption (dummy)									-0,3099 (0,2159)	
Log (No. of app purchases + 1)									0,2715* (0,1213)	0,371*** (0,0826)
Intercept	0,6582*** (71,0371)	-5,2771*** (0,2737)	-4,9075*** (0,6832)	-5,3471*** (0,2866)	-5,1154*** (0,2998)	-5,1412*** (0,2994)	-5,2177*** (0,6883)	-5,1976*** (0,2828)	-5,1367*** (0,3042)	-5,1048*** (0,2816)
R^2 (max scaled)	0,0956	0,0902	0,0963	0,0926	0,0981	0,0961	0,12	0,1143	0,1305	0,1256
Area under the ROC curve	0,726	0,718	0,72	0,713	0,713	0,71	0,752	0,745	0,764	0,762
AIC	937,637	932,599	939,017	930,437	927,441	927,244	913,455	910,657	901,813	902,33
SC	1000,722	964,141	1008,41	961,98	965,292	958,787	970,231	942,199	952,281	940,18
-2 Log L	917,637	922,599	917,017	920,437	915,441	917,244	895,455	900,657	885,813	890,33
N	4058	4058	4058	4058	4058	4058	4058	4058	4058	4058
Events	109	109	109	109	109	109	109	109	109	109

***p<0.001; **p<0.01; *p<0.05; 'p<0.10.

Table 3: Logistic regression on entrepreneurial action

Predictor variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
COMMUNICATION										
Demographics										
US residence (dummy)	0,7175 (0,7473)	0,4704 (0,7052)	0,7895 (0,7885)	0,4979 (0,7239)	0,5656 (0,7396)	0,4979 (0,7239)	0,0237 (0,8067)	0,3904 (0,7401)	0,647 (0,8012)	0,647 (0,8012)
Technical ability proxy	0,6297 (1,1809)	0,226 (1,2266)	1,3489 (1,3754)	0,7788 (1,3349)	0,8669 (1,3453)	0,7788 (1,3349)	0,8576 (1,4074)	0,1282 (1,1601)	-4,1396' (2,5063)	-4,1396' (2,5063)
Communication metrics										
Log (No. of posts)	-0,9881' (0,5887)									
Log (No. of threads started + 1)	0,663 (0,5906)									
Average length of posts	0,000193 (0,00234)									
Extrovert conversation partners	-0,00097 (0,00152)									
Sentiment coding (positive tone)	0,0672 (2,1196)									
Encouragement (inbound sentiment coding)	8,342** (2,7119)	5,269** (1,7964)	8,7261** (2,7525)	5,8017** (1,8608)	6,2358** (2,0785)	5,8017** (1,8608)	8,0833*** (2,3457)	6,6119*** (1,9803)	7,7384** (2,384)	7,7384** (2,384)
Attentiveness proxy	-0,00289 (0,8029)									
Social network metrics										
Forum tenure			-0,0164 (0,0115)	-0,0183' (0,0111)	-0,0173 (0,0109)	-0,0183' (0,0111)	-0,0184 (0,0122)			
Log (No. of unique contacts + 1)			-0,00168 (0,3376)							
No. of fora frequented			-0,4097 (0,2849)							
Log (Degree centrality)			0,3214 (0,5068)							
Prestige			-0,2415 (2,9561)							
Burt's measure of constraint			3,1405 (2,2196)							
Contagion										
Log (In-degree from entrepreneurs +1)					-0,1403 (0,2926)					
CONSUMPTION										
Early adopter metrics										
Product tenure (base product)							-0,1012 (0,1821)			
Avg. time to purchase of apps							0,00288 (0,00531)			
Time to first repurchase of apps							0,00859 (0,00587)	0,00411** (0,00155)	0,0172* (0,00678)	0,0172* (0,00678)
Pct. apps from third-party developers							3,6871 (3,3432)			
Bulk consumption metrics										
No. of base product purchased									1,0385* (0,5012)	1,0385* (0,5012)
Old app-type consumption (dummy)									-2,144* (1,0796)	-2,144* (1,0796)
Log (No. of app purchases + 1)									2,4559* (1,1075)	2,4559* (1,1075)
Intercept	-3,4057 (2,2013)	-4,8198*** (1,0597)	-6,2778* (2,5544)	-4,5258*** (1,1237)	-4,627*** (1,1549)	-4,5258*** (1,1237)	-9,0392* (3,7081)	-6,1775*** (1,2371)	-12,8572** (3,9113)	-12,8572** (3,9113)
R^2 (max scaled)	0,2496	0,1783	0,2895	0,2247	0,228	0,2247	0,3489	0,2845	0,4168	0,4168
Area under the ROC curve	0,898	0,822	0,906	0,871	0,87	0,871	0,901	0,863	0,899	0,899
AIC	83,625	76,816	80,671	75,45	77,212	75,45	74,195	71,044	66,967	66,967
SC	117,516	90,372	114,562	92,396	97,547	92,396	104,697	87,989	94,08	94,08
-2 Log L	63,625	68,816	60,671	65,45	65,212	65,45	56,195	61,044	50,967	50,967
N	219	219	219	219	219	219	219	219	219	219
Events	10	10	10	10	10	10	10	10	10	10

***$p<0.001$; **$p<0.01$; *$p<0.05$; '$p<0.10$.

Table 4: Univariate t-tests between nascent entrepreneurs and users

Measure	$\bar{x}_{nascent\ entrepreneurs} - \bar{x}_{users}$	t-value	p-value	N
COMMUNICATION				
Communication metrics				
Log (No. of posts)	1,0903	11,56	<.0001	4279
Log (No. of threads started + 1)	0,4766	8,01	<.0001	4279
Average length of posts	-103,1529	-6.00	<.0001	4279
Extrovert conversation partners	63,724	2,42	0,0154	4279
Sentiment coding (positive tone)	0,0564	4,29	<.0001	4279
Encouragement (inbound sentiment coding)	0,0768	6,71	<.0001	4279
Attentiveness proxy	-0,0128	-0,22	0,8234	4279
Social network metrics				
Forum tenure	20,1613	7,99	<.0001	4279
Log (No. of unique contacts + 1)	0,5353	7,90	<.0001	4279
No. of fora frequented	1,2805	13,55	<.0001	4279
Log (Degree centrality)	1,0919	11,15	<.0001	4279
Prestige	-0,0239	-1,62	0,1055	4279
Burt's measure of constraint	-0,1557	-7,24	<.0001	4279

Table 5: Summary of predictors of entrepreneurial progression

Predictors	Intent	Action
COMMUNICATION		
Demographics		
Technical ability proxy	(+)	
Communication metrics		
Log (No. of posts)	(+)	
Log (No. of threads started + 1)	(-)	
Encouragement (inbound sentiment coding)		+
Social network metrics		
Forum tenure		(-)
Log (Degree centrality)	(+)	
Contagion		
Log (In-degree from entrepreneurs +1)	+	
CONSUMPTION		
Early adopter metrics		
Time to first repurchase of apps		+
Pct. apps from third-party developers	(+)	
Bulk consumption metrics		
No. of base product purchased	+	+
Old app-type consumption (dummy)		-
Log (No. of app purchases + 1)	+	+